AD-A060 717 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
A RECURSIVE APPROACH TO PARAMETER ESTIMATION IN REGRESSION AND --ETC(U)
MAY 78 J LEDOLTER DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1853 NL

1 OF 1
AD
A060717

END
DATE
FILMED
01-79
DDC

MRC Technical Summary Report #1853

A RECURSIVE APPROACH TO PARAMETER ESTIMATION IN REGRESSION AND TIME SERIES MODELS

Johannes Ledolter

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

May 1978

(Received March 29, 1978)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

6

A RECURSIVE APPROACH TO PARAMETER ESTIMATION IN REGRESSION AND TIME SERIES MODELS†

10 Johannes Ledolter

11 May 78

9 Technical Summary Report, #1853
~~May 1978~~

ABSTRACT

14 MRC-TSR-1853

This paper discusses the recursive (on line) estimation of parameters in regression and autoregressive integrated moving average (ARIMA) time series models. The approach which is adopted uses Kalman filtering techniques to calculate estimates recursively. This approach can be used for the case of constant as well as time varying parameters.

In the first section the linear regression model is considered and recursive estimates of the parameters, both for constant and time varying parameters, are discussed. Since the stochastic model for the parameters over time will be rarely known, simplifying assumptions have to be made. In particular a random walk as a model for time varying parameters is assumed and it is shown how one can determine whether the parameters are constant or changing over time.

In the second section the recursive estimation of parameters in ARIMA models is considered. If moving average terms are present, the model has to be linearized and the Extended Kalman Filter can be used to recursively update the parameter estimates. The first order moving average model is discussed in detail.

15 DAAG29-75-C-0024, NSF-MCS75-17385

AMS (MOS) Subject Classifications: 62M10, 62J05

Key Words: Regression, ARIMA time series model, Recursive estimation, Kalman Filter, Extended Kalman Filter

Work Unit Number 4 (Probability, Statistics, and Combinatorics)

12 18p.

† Paper presented at the IMS Special Topics Meeting on Time Series Analysis, Ames, Iowa, May 1-3, 1978.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Science Foundation under Grant No. MCS75-17385 A01.

221 200

SIGNIFICANCE AND EXPLANATION

The problem discussed in this paper is the following: Observations up to time n (of, for example, business data or missile positions) are available and one wants to estimate unknown parameters in regression or autoregressive integrated moving average models in order to predict future values.

An additional observation is recorded. The question becomes how to update the parameter estimates from the previous estimates and the most recent observation without storing the complete past history of the data. The answer to this question will depend on whether the parameters in these models are assumed constant or whether they themselves follow a given stochastic process.

This recursive estimation procedure (which is sometimes called on-line estimation or parameter tracking) is important if the observations become available sequentially in time. It has applications in economics and business, where economic indicators and sales data are updated every week, month or quarter; it can also be applied to the control of satellites or ballistic missiles where the position in space is recorded every few seconds.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

A RECURSIVE APPROACH TO PARAMETER ESTIMATION IN REGRESSION AND TIME SERIES MODELS†

Johannes Ledolter

1. Recursive parameter estimation in regression models

§1.1: We consider n observations from the regression model

$$z_t = \underset{\sim}{h}_t' \underset{\sim}{\beta}_t + a_t \tag{1.1}$$

where: z_t is the dependent variable at time t

$\underset{\sim}{h}_t = (h_{1t}, \ldots, h_{mt})$ is a set of m independent variables at time t

$\{a_t\}$ is a sequence of independent normal random variables $N(0, \sigma^2)$.

Standard regression procedures assume that the parameters are constant over time ($\underset{\sim}{\beta}_t = \underset{\sim}{\beta}$ for all t). In this case the least squares (maximum likelihood) estimates of $\underset{\sim}{\beta}$, given all the data up to and including time n, are

$$\hat{\underset{\sim}{\beta}}_n = (H^{(n)'} H^{(n)})^{-1} H^{(n)'} \underset{\sim}{z}^{(n)}$$

where

$$\underset{\sim}{z}^{(n)} = (z_1, z_2, \ldots, z_n)'$$

$$H^{(n)} = (\underset{\sim}{h}_1, \underset{\sim}{h}_2, \ldots, \underset{\sim}{h}_n)' .$$

A recursive version of least squares estimates in regression models was first given by Plackett [8]. He proves that the parameter estimates at time n are linear combinations of the estimates at time $n - 1$ and the prediction error $z_n - \underset{\sim}{h}_n' \hat{\underset{\sim}{\beta}}_{n-1}$.

Plackett's solution can be shown to be a special case of a more general procedure which, in the engineering literature, is known under the name of Kalman filtering (Kalman [4], Kalman and Bucy [5]). In the following section (§1.2) we given a brief review of this approach. A similar discussion is given by Duncan and Horn [2].

---

†Paper presented at the IMS Special Topics Meeting on Time Series Analysis, Ames, Iowa, May 1-3, 1978.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Science Foundation under Grant No. MCS75-17385 A01.

§1.2: We consider the regression model in (1.1)

$$z_t = \underset{\sim}{h}_t' \underset{\sim}{\beta}_t + a_t \tag{1.2}$$

with parameters which change according to a multiple (m-dimensional) autoregressive model

$$\underset{\sim}{\beta}_{t+1} = T\underset{\sim}{\beta}_t + \underset{\sim}{\epsilon}_{t+1} \tag{1.3}$$

$\{\underset{\sim}{\epsilon}_t\}$ is a sequence of independent normal random variables $N_m(\underset{\sim}{0},\sigma^2\Omega)$. Furthermore it is assumed that $\underset{\sim}{\epsilon}_t$ and a_t are independent; $Ea_t\underset{\sim}{\epsilon}_t = \underset{\sim}{0}$.

Assuming that $\underset{\sim}{\beta}_0$, the initial parameter vector at time 0, follows a normal distribution with expectation $\underset{\sim}{\beta}_{0|0}$ and covariance matrix $\sigma^2 P_{0|0}$, it is easily shown that:

$$p(\underset{\sim}{\beta}_t, z_t | \underset{\sim}{z}^{(t-1)}, H^{(t)}, \sigma^2, \Omega, T, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \sim N_{m+1}\left\{\begin{bmatrix} \hat{\underset{\sim}{\beta}}_{t|t-1} \\ \underset{\sim}{h}_t'\hat{\underset{\sim}{\beta}}_{t|t-1} \end{bmatrix}, \sigma^2 \begin{bmatrix} P_{t|t-1} & P_{t|t-1}\underset{\sim}{h}_t \\ \underset{\sim}{h}_t' P_{t|t-1} & 1+\underset{\sim}{h}_t' P_{t|t-1}\underset{\sim}{h}_t \end{bmatrix}\right\} \tag{1.4}$$

and

$$p(\underset{\sim}{\beta}_t | \underset{\sim}{z}^{(t)}, H^{(t)}, \sigma^2, \Omega, T, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \sim N_m\{\hat{\underset{\sim}{\beta}}_{t|t}, \sigma^2 P_{t|t}\} . \tag{1.5}$$

It follows from properties of the normal distribution that

$$\hat{\underset{\sim}{\beta}}_{t|t} = \hat{\underset{\sim}{\beta}}_{t|t-1} + K_{t-1}(z_t - \underset{\sim}{h}_t'\hat{\underset{\sim}{\beta}}_{t|t-1}) \tag{1.6a}$$

$$P_{t|t} = P_{t|t-1} - K_{t-1}\underset{\sim}{h}_t' P_{t|t-1} . \tag{1.6b}$$

$$K_{t-1} = P_{t|t-1}\underset{\sim}{h}_t'(1 + \underset{\sim}{h}_t' P_{t|t-1}\underset{\sim}{h}_t)^{-1} \tag{1.6c}$$

is called Kalman gain. *Furthermore it follows from (1.4) that*

$$p(\underset{\sim}{\beta}_{t+1} | \underset{\sim}{z}^{(t)}, H^{(t)}, \sigma^2, \Omega, T, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \sim N_m(\hat{\underset{\sim}{\beta}}_{t+1|t}, \sigma^2 P_{t+1|t})$$

where

$$\hat{\underset{\sim}{\beta}}_{t+1|t} = T\hat{\underset{\sim}{\beta}}_{t|t} \tag{1.6d}$$

$$P_{t+1|t} = TP_{t|t}T' + \Omega . \tag{1.6e}$$

Equations (1.6a)-(1.6e) can be used to calculate the parameter estimates recursively as each new observation becomes available in time. $\underset{\sim}{\beta}_{0|0}$ and $P_{0|0}$ are starting values for the recursive estimation procedure in equations (1.6). To reflect ignorance about the parameters at time zero (before observations become available), the matrix $P_{0|0}$ is usually chosen diagonal with large values in its diagonal. If n is moderately large, the initial choice of $\underset{\sim}{\beta}_{0|0}$ will be dominated by the information from the data.

§1.3: For the case $T = I$ and $\Omega = 0$ the above procedure specializes to the recursive least squares algorithm discussed by Plackett [8].

$$\hat{\underset{\sim}{\beta}}_{t|t} = \hat{\underset{\sim}{\beta}}_{t-1|t-1} + (1 + \underset{\sim}{h}_t' P_{t-1|t-1}\underset{\sim}{h}_t)^{-1} P_{t-1|t-1}\underset{\sim}{h}_t(z_t - \underset{\sim}{h}_t'\hat{\underset{\sim}{\beta}}_{t-1|t-1}) \tag{1.7a}$$

and

$$P_{t|t} = P_{t-1|t-1} - (1 + \underset{\sim}{h}_t' P_{t-1|t-1}\underset{\sim}{h}_t)^{-1} P_{t-1|t-1}\underset{\sim}{h}_t\underset{\sim}{h}_t' P_{t-1|t-1} \,. \tag{1.7b}$$

§1.4: Equation (1.3) specifies an autoregressive process as model for the time varying parameters. This, however, is not a serious restriction since by the introduction of additional state variables any autoregressive integrated moving average (ARIMA) process can be brought into this form.

By substitution it can be shown that the regression model (1.2) in which the parameters $\underset{\sim}{\beta}_t$ follow a multiple ARIMA process

$$(I - \Phi_1 B - \cdots - \Phi_p B^p)\underset{\sim}{\beta}_{t+1} = (I - \Theta_1 B - \cdots - \Theta_q B^q)\underset{\sim}{\epsilon}_{t+1} \tag{1.8}$$

where I is the identity matrix and $\Phi_1,\ldots,\Phi_p,\Theta_1,\ldots,\Theta_q$ are known $[m \times m]$ matrices, can be written as:

$$\begin{bmatrix} \underset{\sim}{\beta}_{t+1} \\ \underset{\sim}{\beta}^*_{2t+1} \\ \vdots \\ \underset{\sim}{\beta}^*_{kt+1} \end{bmatrix} = \left[\begin{array}{c|c} \Phi_1 & \\ \vdots & I_{(k-1)m} \\ \hline \Phi_k & 0 \cdots 0 \end{array}\right] \begin{bmatrix} \underset{\sim}{\beta}_t \\ \underset{\sim}{\beta}^*_{2t} \\ \vdots \\ \underset{\sim}{\beta}^*_{kt} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{\epsilon}_{t+1} \\ \underset{\sim}{\epsilon}^*_{2t+1} \\ \vdots \\ \epsilon^*_{kt+1} \end{bmatrix} \tag{1.9}$$

and

$$z_t = [\underset{\sim}{h}_t' \underset{\sim}{0}' \cdots \underset{\sim}{0}'] \begin{bmatrix} \underset{\sim}{\beta}_t \\ \underset{\sim}{\beta}^*_{2t} \\ \vdots \\ \underset{\sim}{\beta}^*_{kt} \end{bmatrix} + a_t \tag{1.10}$$

$k = \max\{p, q+1\}$; $\Phi_j = 0$ for $j > p$; $\Theta_j = 0$ for $j > q$; $I_{(k-1)m}$ is the $(k-1)m$ identity matrix; $\epsilon^*_{j\,t+1} = -\Theta_{j-1}\epsilon_{t+1}$ for $2 \le j \le k$.

The recursive procedures given in (1.6) can be applied to the system described by equations (1.9) and (1.10).

§1.5: The model in equation (1.4) assumes that the parameters β_t vary around the origin. In general, however, parameters will vary around a level $\bar{\beta}$ which will be different from zero.

Redefining the matrix T will, however, allow for $\bar{\beta} \neq 0$:

$$\begin{bmatrix} \beta_{t+1} \\ \bar{\beta}_{t+1} \end{bmatrix} = \begin{bmatrix} T & T - I \\ 0 & I \end{bmatrix} \begin{bmatrix} \beta_t \\ \bar{\beta}_t \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} \epsilon_{t+1} \qquad (1.11)$$

$$z_t = [h_t' 0'] \begin{bmatrix} \beta_t \\ \bar{\beta}_t \end{bmatrix} + a_t . \qquad (1.12)$$

§1.6: A random walk model for the parameters

The recursive estimation procedures in equations (1.6) are very general. Theoretically, they can be used to update regression estimates for any given ARIMA model for the parameters.

In practice, however, the model for the parameters is rarely known and simplifying assumptions about this model have to be made. A particularly useful and simple model for time varying parameters is the random walk model,

$$\beta_{t+1} = \beta_t + \epsilon_{t+1} \qquad (1.13)$$

which equivalently can be written as

$$\beta_{t+1} = \beta_0 + \sum_{j=0}^{t} \epsilon_{t+1-j} .$$

It represents the parameter at time $t + 1$ as a sum of the parameter at time 0 and a cumulative sum of independent random variables. The random walk model allows for smooth changes in the parameters. The variability of the parameters depends on the covariance matrix $E\epsilon_t\epsilon_t' = \sigma^2\Omega$. If $\Omega = 0$ the parameters are constant over time.

Using the observations up to and including time n, we address the question whether the parameters in the regression model (1.2) can be considered constant, or whether there is indication of time changing parameters ($\Omega \neq 0$ in the random walk model).

Setting $T = I$, it follows from equation (1.4):

$$p(\underset{\sim}{z}_t | \underset{\sim}{z}^{(t-1)}, H^{(t)}, \sigma^2, \Omega, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \sim N(\underset{\sim}{h}_t' \hat{\underset{\sim}{\beta}}_{t|t-1}; \sigma^2 f_t) \tag{1.14}$$

where

$$f_t = 1 + \underset{\sim}{h}_t' P_{t|t-1} \underset{\sim}{h}_t .$$

Furthermore,

$$\begin{aligned} p(z_1, \ldots, z_n | H^{(n)}, \sigma^2, \Omega, \underset{\sim}{\beta}_{0|0}, P_{0|0}) &= p(z_1 | H^{(1)}, \sigma^2, \Omega, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \times \\ &\times \prod_{t=2}^{n} p(z_t | \underset{\sim}{z}^{(t-1)}, H^{(t)}, \sigma^2, \Omega, \underset{\sim}{\beta}_{0|0}, P_{0|0}) \\ &\propto \sigma^{-n} \left[\prod_{t=1}^{n} f_t \right]^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^{n} \frac{(z_t - \underset{\sim}{h}_t' \hat{\underset{\sim}{\beta}}_{t|t-1})^2}{f_t} \right\} . \end{aligned} \tag{1.15}$$

$\underset{\sim}{\beta}_{0|0}$ and $P_{0|0}$ can be treated as parameters. However, if there is little prior information about $\underset{\sim}{\beta}_0$, the matrix $P_{0|0}$ in the recursions (1.6) can be chosen diagonal with large entries in its diagonal; for the initial $\underset{\sim}{\beta}_{0|0}$ one can choose the zero vector.

The log likelihood function of the parameters σ^2 and Ω is given by

$$\ell(\sigma^2, \Omega | \underset{\sim}{z}^{(n)}, H^{(n)}) \propto -n \log \sigma - \frac{1}{2} \sum_{t=1}^{n} \log f_t - \frac{1}{2\sigma^2} \sum_{t=1}^{n} \frac{(z_t - \underset{\sim}{h}_t' \hat{\underset{\sim}{\beta}}_{t|t-1})^2}{f_t} . \tag{1.16}$$

The ML estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \frac{(z_t - \underset{\sim}{h}_t' \hat{\underset{\sim}{\beta}}_{t|t-1})^2}{f_t} \tag{1.17}$$

and the concentrated likelihood

$$\ell_c(\Omega | \underset{\sim}{z}^{(n)}, H^{(n)}) \propto -n \log \hat{\sigma} - \frac{1}{2} \sum_{t=1}^{n} \log f_t . \tag{1.18}$$

Nonlinear optimization techniques can be used to derive the ML estimate of the covariance matrix Ω. Approximate $(1 - \alpha)100$ percent confidence limits for the $\frac{k(k+1)}{2}$ elements in Ω are given by the contours of

$$|\ell_c(\hat{\Omega}|z^{(n)}, H^{(n)}) - \ell_c(\Omega|z^{(n)}, H^{(n)})| \le \frac{1}{2}\chi^2_{\frac{k(k+1)}{2}}(\alpha)$$

where $\chi^2_f(\alpha)$ is the upper α percent cut off value of a χ^2 distribution with f degrees of freedom.

§1.7: An example

We consider the model

$$z_t = \beta_t + a_t \qquad V(a_t) = \sigma^2 \tag{1.19a}$$

$$\beta_{t+1} = \beta_t + \epsilon_{t+1} \qquad V(\epsilon_t) = k\sigma^2 \qquad k \ge 0 . \tag{1.19b}$$

(i) The updating relations (1.6) simplify to

$$\hat{\beta}_{t|t} = \hat{\beta}_{t-1|t-1} + (P_{t-1|t-1} + k + 1)^{-1}(P_{t-1|t-1} + k)(z_t - \hat{\beta}_{t-1|t-1}) \tag{1.20a}$$

$$P_{t|t} = (P_{t-1|t-1} + k)/(P_{t-1|t-1} + k + 1) \tag{1.20b}$$

and the concentrated log likelihood function in (1.18)

$$\ell_c(k|z^{(n)}) \propto -n \log \hat{\sigma} - \frac{1}{2}\sum_{t=1}^{n} \log f_t \tag{1.21}$$

where

$$f_t = P_{t-1|t-1} + k + 1$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} \frac{(z_t - \hat{\beta}_{t-1|t-1})^2}{P_{t-1|t-1} + k + 1} .$$

The log likelihood function can be calculated and plotted for various k. The maximum likelihood estimate $\hat{k}$ can be found and approximate $(1 - \alpha)100\%$ confidence limits are given by the solution to

$$|\ell_c(\hat{k}|z^{(n)}) - \ell_c(k|z^{(n)})| = \frac{1}{2}\chi^2_1(\alpha) . \tag{1.22}$$

(ii) The model in (1.19) can equivalently be represented as integrated first order moving average process (Muth [7])

$$z_t = z_{t-1} + \xi_t - \theta\xi_{t-1} \tag{1.23}$$

where

$$\theta = \frac{1}{2}\,[(2 + k) - \sqrt{4k + k^2}], \quad \theta \geq 0$$

$$\sigma_\xi^2 = \frac{2\sigma^2}{(2 + k) - \sqrt{4k + k^2}} .$$

This equivalent representation can be used as check whether the model in equations (1.19) is justified by the data.

2. Recursive estimation of parameters in ARIMA time series models

§2.1: Autoregressive integrated moving average (ARIMA) time series models are described by the stochastic difference equation

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(z_t - \mu) = (1 - \theta_1 B - \cdots - \theta_q B^q) a_t \tag{2.1}$$

where: B is the backshift operator; $B^m z_t = z_{t-m}$

z_t is a stationary difference of the original observations, $z_t = (1 - B)^d x_t$

the roots of $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$ and

$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q = 0$ are assumed outside the unit circle;

furthermore $\phi(B) = 0$ and $\theta(B) = 0$ have no common roots.

$\{a_t\}$ is a sequence of independent $N(0,\sigma^2)$ random variables.

§2.2: Given observations $\underset{\sim}{z}^{(n)} = (z_1, z_2, \ldots, z_n)'$, maximum likelihood estimates for the parameters $\underset{\sim}{\beta}' = (\mu, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ can be derived. Box and Jenkins [1], Ljung and Box [6] discuss the derivation of the exact likelihood function of the parameters for this class of models. In general, the exact likelihood function is nonlinear in the parameters and iterative maximization techniques have to be used for the derivation of the ML estimates. Ljung and Box [6] propose a general method for the calculation of the likelihood function.

For pure autoregressive processes the derivation can be simplified. The ML estimates of $\underset{\sim}{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)'$ can be approximated by the least squares estimates.

§2.3: ML estimates for the parameters in ARIMA models are nonrecursive in nature. If data is collected sequentially (such as quarterly or monthly economic data, sales data, hourly pollutant measurements) one would, however, prefer recursive estimation procedures which update the values of the parameters as each new observation becomes available.

Updating (tracking) the parameters should not be confused with updating the forecasts as discussed by Box and Jenkins [1, Chapter 5]. Updating of the forecasts assumes constant parameters and calculates the revised forecast as each new observation is made available. Tracking the parameters updates the parameter estimates.

§2.4: Linearization of the model

The ARIMA model in (2.1) is nonlinear in the parameters. To derive recursive estimates the model has to be linearized. This can be achieved by expanding the ARIMA model

$$z_t = \mu + \sum_{j\geq 1} \pi_j(\underset{\sim}{\phi},\underset{\sim}{\theta})(z_{t-j} - \mu) + a_t \tag{2.2}$$

where the π_j weights in (2.2) are given by the coefficients in

$$\pi(B) = 1 - \sum_{j\geq 1} \pi_j B^j = \frac{\phi(B)}{\theta(B)} \tag{2.3}$$

in a Taylor series around some reference value $\bar{\underset{\sim}{\beta}} = (\bar{\mu},\bar{\underset{\sim}{\phi}},\bar{\underset{\sim}{\theta}})'$.

Define

$$f(\underset{\sim}{\beta};\underset{\sim}{z}^{(t-1)}) = \mu + \sum_{j\geq 1} \pi_j(\underset{\sim}{\phi},\underset{\sim}{\theta})(z_{t-j} - \mu) ; \tag{2.4}$$

then,

$$z_t \simeq f(\bar{\underset{\sim}{\beta}},\underset{\sim}{z}^{(t-1)}) + (\mu - \bar{\mu})u_t(\bar{\underset{\sim}{\beta}}) + \sum_{i=1}^{p} (\phi_i - \bar{\phi}_i)v_{it}(\bar{\underset{\sim}{\beta}}) + \sum_{j=1}^{q} (\theta_j - \bar{\theta}_j)w_{jt}(\bar{\underset{\sim}{\beta}}) + a_t \tag{2.5}$$

where

$$u_t(\bar{\underset{\sim}{\beta}}) = \frac{d}{d\mu} f(\underset{\sim}{\beta};\underset{\sim}{z}^{(t-1)})\Big|_{\underset{\sim}{\beta}=\bar{\underset{\sim}{\beta}}}$$

$$v_{it}(\bar{\underset{\sim}{\beta}}) = \frac{d}{d\phi_i} f(\underset{\sim}{\beta};\underset{\sim}{z}^{(t-1)})\Big|_{\underset{\sim}{\beta}=\bar{\underset{\sim}{\beta}}} \qquad (1 \leq i \leq p)$$

$$w_{jt}(\bar{\underset{\sim}{\beta}}) = \frac{d}{d\theta_j} f(\underset{\sim}{\beta};\underset{\sim}{z}^{(t-1)})\Big|_{\underset{\sim}{\beta}=\bar{\underset{\sim}{\beta}}} \qquad (1 \leq j \leq q) . \tag{2.6}$$

$z_t - f(\bar{\underset{\sim}{\beta}};z^{(t-1)}) = a_t(\bar{\underset{\sim}{\beta}})$ *is the one step ahead forecast error for the* ARIMA model with parameters $\bar{\underset{\sim}{\beta}}$.

It can be easily shown that

$$u_t(\bar{\underset{\sim}{\beta}}) = (1 - \bar{\phi}_1 - \cdots - \bar{\phi}_p)/(1 - \bar{\theta}_1 - \cdots - \bar{\theta}_q)$$

$$(1 - \bar{\phi}_1 B - \cdots - \bar{\phi}_p B^p)v_{it}(\bar{\underset{\sim}{\beta}}) = a_{t-i}(\bar{\underset{\sim}{\beta}})$$

$$(1 - \bar{\theta}_1 B - \cdots - \bar{\theta}_q B^q)w_{jt}(\bar{\underset{\sim}{\beta}}) = -a_{t-j}(\bar{\underset{\sim}{\beta}}) . \tag{2.7}$$

Using this expansion the general ARIMA model (2.1) with time varying coefficients

$$\beta_{t+1} = T\beta_t + \epsilon_{t+1}; \quad E\epsilon_t\epsilon_t' = \sigma^2\Omega \tag{2.8}$$

can be expanded around the trajectory $\bar{\beta}_{t+1} = T\bar{\beta}_t$ and represented as

$$\begin{cases} z_t - f(\bar{\beta}_t; z^{(t-1)}) \simeq h'(\bar{\beta}_t; z^{(t-1)})(\beta_t - \bar{\beta}_t) + a_t \\ (\beta_{t+1} - \bar{\beta}_{t+1}) = T(\beta_t - \bar{\beta}_t) + \epsilon_{t+1} . \end{cases} \tag{2.9}$$

The elements in $h(\bar{\beta}_t, z^{(t-1)}) = (u_t, v_{1t}, \ldots, v_{pt}, w_{1t}, \ldots, w_{qt})'$ are given by the expressions in (2.7) with $\bar{\beta}$ replaced by $\bar{\beta}_t$.

Summarizing, it is shown that the nonlinear ARIMA model with time varying parameters (the special case of constant parameters is given when $T = I$ and $\Omega = 0$) can be linearized and approximated as in (2.9). (2.9) is linear in the measurement deviations $z_t - f(\bar{\beta}_t; z^{(t-1)})$ and in the trajectory (state) deviations $\beta_t - \bar{\beta}_t$.

Given a trajectory $\bar{\beta}_t$ and observations $z^{(t)}$, one can calculate the measurement deviations $z_t - f(\bar{\beta}_t; z^{(t-1)})$ and use the Kalman filter equations in (1.6) to derive an estimate $(\widehat{d\beta})_{t|t}$ for the deviations from the trajectory $(d\beta)_t = \beta_t - \bar{\beta}_t$. The estimate for β_t is then given by $\hat{\beta}_{t|t} = \bar{\beta}_t + (\widehat{d\beta})_{t|t}$.

An obvious choice for the reference trajectory is $\beta_0 = \bar{\beta}_0$, where $\bar{\beta}_0$ is the prior estimate of the parameter vector β. However, if our prior estimate is poor, the state deviations $\beta_t - \bar{\beta}_t$ can become large and the linearity assumption used in the expansion (2.5) may be violated.

However, the following improvement can be made; the ARIMA model (2.9) can be linearized about each new estimate as estimates become available in time. At time t (after z_t has become available) we linearize around $\hat{\beta}_{t|t}$. Processing z_{t+1} we can update the parameter vector, derive $\hat{\beta}_{t+1|t+1}$ and use this estimate for the next linearization. This procedure uses a better estimate for the trajectory as soon as one becomes available, and large errors about the trajectory due to bad a priori estimates are not allowed to propagate.

It follows from the recursive linearization of the ARIMA model that

$$(\widehat{d\beta})_{t+1|t} = 0 \tag{2.10}$$

and

$$(\widehat{d\beta})_{t+1|t+1} = \hat{\beta}_{t+1|t+1} - \hat{\beta}_{t+1|t} . \tag{2.11}$$

Using the Kalman filter equations for the linearized system (2.9), the updating equations are therefore given by:

$$\hat{\beta}_{t+1|t} = T\hat{\beta}_{t|t} , \tag{2.12a}$$

$$P_{t+1|t} = TP_{t|t}T' + \Omega , \tag{2.12b}$$

$$\hat{\beta}_{t|t} = \hat{\beta}_{t|t-1} + K_{t-1}(z_t - f(\hat{\beta}_{t|t-1};z^{(t-1)})) , \tag{2.12c}$$

$$P_{t|t} = P_{t|t-1} - K_{t-1}h'(\hat{\beta}_{t|t-1};z^{(t-1)})P_{t|t-1} , \tag{2.12d}$$

$$K_{t-1} = P_{t|t-1}h'(\hat{\beta}_{t|t-1},z^{(t-1)})[1 + h'(\hat{\beta}_{t|t-1},z^{(t-1)})P_{t|t-1}h(\hat{\beta}_{t|t-1},z^{(t-1)})]^{-1} . \tag{2.12e}$$

In the engineering literature the general procedure of recursively linearizing non-linear models is known under the name Extended Kalman Filtering. See Jazwinski [3] for further reference.

§2.5: Example: Recursive parameter estimation for the first order moving average process with constant parameter

(i) The first order moving average process (for $\mu = 0$) is given by

$$z_t = a_t - \theta a_{t-1} . \tag{2.13}$$

It is, for example, shown by Ljung and Box [6] that the ML estimate of θ for the first order moving average model is obtained by minimizing

$$L(\theta) = \left\{ \sum_{j=0}^{n} \theta^{2j} \right\}^{\frac{1}{n}} S(\theta) \tag{2.14}$$

where

$$S(\theta) = \sum_{t=0}^{n} \hat{a}_t^2(\theta) ,$$

$$\hat{a}_0(\theta) = -\left[\sum_{j=0}^{n} \theta^{2j}\right]^{-1} \sum_{i=1}^{n} \theta^i \left(1 + \sum_{j=1}^{n+1-i} \theta^{2(n-j)}\right) z_i$$

and

$$\hat{a}_t(\theta) = z_t + \theta \hat{a}_{t-1}(\theta) \qquad (1 \leq t \leq n) .$$

(ii) To derive parameter estimates recursively we use the theory outlined above. For the case of a first order moving average process with constant parameter θ, equations (2.12) simplify to (for notational convenience we write $\hat{\theta}_t = \hat{\theta}_{t|t}$ and $P_t = P_{t|t}$)

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{P_{t-1} h(\hat{\theta}_{t-1}; \underset{\sim}{z}^{(t-1)})}{P_{t-1} h(\hat{\theta}_{t-1}; \underset{\sim}{z}^{(t-1)})^2 + 1} [z_t + \hat{\theta}_{t-1} a_{t-1}(\hat{\theta}_{t-1})] , \tag{2.15a}$$

$$P_t = P_{t-1} / (P_{t-1} h(\hat{\theta}_{t-1}; \underset{\sim}{z}^{(t-1)})^2 + 1) \tag{2.15b}$$

where

$$h(\hat{\theta}_{t-1}; \underset{\sim}{z}^{(t-1)}) = w_t(\hat{\theta}_{t-1})$$

which according to (2.7) can be updated by

$$w_t(\hat{\theta}_{t-1}) = \hat{\theta}_{t-1} w_{t-1}(\hat{\theta}_{t-2}) - a_{t-1}(\hat{\theta}_{t-1}) \tag{2.16}$$

and

$$a_t(\hat{\theta}_t) = z_t + \hat{\theta}_t a_{t-1}(\hat{\theta}_{t-1}) .$$

REFERENCES

[1] Box, G. E. P. and G. M. Jenkins (1970): Time Series Analysis, Forecasting and Control; Holden Day, San Francisco.

[2] Duncan, D. B. and S. D. Horn (1972): Linear dynamic regression estimation from the viewpoint of regression analysis; JASA, 67, 815-821.

[3] Jazwinski, A. H. (1970): Stochastic Processes and Filtering Theory; Academic Press, New York.

[4] Kalman, R. E. (1960): A new approach to linear filtering and prediction problems; Journal of Basic Engineering, 82, 35-45.

[5] Kalman, R. E. and R. S. Bucy (1961): New results in linear filtering and prediction theory; Journal of Basic Engineering, 83, 95-108.

[6] Ljung, G. M. and G. E. P. Box (1976): Studies in the modelling of discrete time series - II. Maximum likelihood estimation in the autoregressive-moving average model; Technical Report 476, Department of Statistics, University of Wisconsin-Madison, October 1976.

[7] Muth, G. F. (1960): Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components; JASA, 55, 299-306.

[8] Plackett, R. L. (1950): Some theorems in least squares; Biometrika, 37, 149-157.

JL/scr

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 1853	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A RECURSIVE APPROACH TO PARAMETER ESTIMATION IN REGRESSION AND TIME SERIES MODELS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Johannes Ledolter		8. CONTRACT OR GRANT NUMBER(s) DAAG29-75-C-0024 MCS75-17385 A01
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Probability, Statistics, and Combinatorics
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below.		12. REPORT DATE May 1978
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Regression
ARIMA time series model
Recursive estimation
Kalman Filter
Extended Kalman Filter

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper discusses the recursive (on line) estimation of parameters in regression and autoregressive integrated moving average (ARIMA) time series models. The approach which is adopted uses Kalman filtering techniques to calculate estimates recursively. This approach can be used for the case of constant as well as time varying parameters. → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Cont'd.)

In the first section the linear regression model is considered and recursive estimates of the parameters, both for constant and time varying parameters, are discussed. Since the stochastic model for the parameters over time will be rarely known, simplifying assumptions have to be made. In particular a random walk as a model for time varying parameters is assumed and it is shown how one can determine whether the parameters are constant or changing over time.

In the second section the recursive estimation of parameters in ARIMA models is considered. If moving average terms are present, the model has to be linearized and the Extended Kalman Filter can be used to recursively update the parameter estimates. The first order moving average model is discussed in detail.